Novel cytochrome P450 monooxygenases and their use for the oxidation of organic compounds

The present invention relates to novel cytochrome P450 monooxygenases with modified substrate specificity which are capable of the oxidation of organic substrates, for example N-heterocyclic aromatic compounds, nucleotide sequences coding therefor, expression constructs and vectors comprising these sequences, microorganisms transformed therewith, processes for the microbiological oxidation of various organic substrates, such as N-heterocyclic aromatic compounds and in particular processes for the preparation of indigo and indirubin.

Enzymes having novel functions and properties can be prepared either by screening of natural samples or by protein engineering of known enzymes. Under certain circumstances, the last-mentioned method can be the more suitable to induce characteristics whose generation by the natural selection route is improbable. Despite numerous attempts at the engineering of enzymes, up to now there are only a few successful studies for promoting the catalytic activity of enzyme mutants with respect to a certain substrate (1-10). In these known cases, the substrates are structurally closely related to the native substrate of the respective enzyme. As yet, there are no reports on the successful engineering of enzymes which, after modification, catalyze the reaction of a compound which structurally is completely different from the native substrate of the enzyme.

The cytochrome P450 monooxygenase isolatable from the bacterium Bacillus megaterium usually catalyzes the subterminal hydroxylation of long-chain, saturated acids and the corresponding amides and alcohols thereof or the epoxidation of unsaturated long-chain fatty acids or saturated fatty acids of medium chain length (11-13). The optimal chain length of saturated fatty acids is 14 to 16 carbon atoms. Fatty acids having a chain length of less than 12 are not hydroxylated (11).

The structure of the heme domain of P450 BM-3 was determined by X-ray structural analysis (14-16). The substrate binding site is present in the form of a long tunnel-like opening which extends from the surface of the molecule as far as the heme molecule and is almost exclusively bordered by hydrophobic amino acid residues. The only charged residues on the surface of the heme domain are the residues Arg47 and Tyr51. It is assumed that these are involved in the binding of the carboxylate group of the substrate by formation of a hydrogen bond (14). The mutation of

Arg47 to Glu brings about inactivation of the enzyme for arachidonic acid (13), but increases its activity compared with C_{12}-C_{14}-alkyltrimethylammonium compounds (17). Substrate utilization for aromatic compounds, in particular mono-, bi- or polynuclear, if desired heterocyclic, aromatics, alkanes, alkenes, cycloalkanes and cycloalkenes, has not been described for this enzyme. Until now, it was therefore assumed in specialist circles that substrates other than the organic substrates hitherto described, for example indole, on account of the clear structural differences from the native substrates of P450 BM-3, in particular on account of the absence of functional groups which could bind to the abovementioned residues in the substrate pocket, are not a substrate.

It is an object of the present invention to make available novel cytochrome P450 monooxygenases having modified substrate specificity or modified substrate profile. In particular, monooxygenase mutants are to be provided which, in comparison with the nonmutated wild-type enzyme, are enzymatically active with structurally clearly different substrates.

Compared to the wild-type enzyme, a "modified substrate profile" can be observed for the mutants according to the invention. In particular, for the mutant in question, an improvement in reactivity is observed, for example an increase of the specific activity (expressed as nmol of converted substrate/minute/nmol of P450 enzyme) and/or of at least one kinetic parameter selected from the group consisting of Kcat, Km and Kcat/Km (for example by at least 1%, such as 10 to 1000%, 10 to 500% or 10 to 100%) in the conversion of at least one of the oxidizable compounds defined in groups a) to d). The oxidation reaction according to the invention comprises the enzyme-catalyzed oxygenation of at least one exogenous (i.e. added to the reaction medium) or endogenous (i.e. already present in the reaction medium) organic substrate. In particular, the oxidation reaction according to the invention comprises a mono- and/or polyhydroxylation, for example a mono- and/or dihydroxylation, at an aliphatic or aromatic C-H group, or an epoxidation at a C=C group which is preferably non-aromatic. Also possible are combinations of the above reactions. Moreover, the immediate reaction product can be converted further in the context of a non-enzymatic subsequent or side reaction. Such combinations of enzymatic and non-enzymatic processes likewise form part of the subject-matter of the present invention.

We have found that the above object is surprisingly achieved by means of novel cytochrome P450 monooxygenases which, for example, are capable of the oxidation of N-heterocyclic bi- or polynuclear aromatic compounds.

In particular, the invention relates to those monooxygenases whose substrate-binding region is capable by means of site-specific mutagenesis of the functional uptake of novel, for example N-heterocyclic substrates.

In a preferred embodiment of the invention, the novel monooxygenases are soluble, i.e. existent in non membrane-bound form, and enzymatically active in this form.

The monooxygenases according to the invention are preferably derived from cytochrome P450 monooxygenases of bacterial origin, as derived, in particular, from cytochrome P450 monooxygenase BM-3 from *Bacillus megaterium* having an amino acid sequence according to SEQ ID NO:2, which has at least one functional mutation, i.e. promoting the oxidation of novel organic substrates (cf. in particular the groups a) to d) of compounds as defined below), for example N-heterocyclic mono-, bi- or polynuclear aromatic compounds, in one of the amino acid sequence regions 172-224 (F/G loop region), 39-43 (ß-strand 1), 48-52 (ß-strand 2), 67-70 (ß-strand 3), 330-335 (ß-strand 5), 352-356 (ß-strand 8), 73-82 (helix 5) and 86-88 (helix 6).

The cytochrome P450 monooxygenase mutants provided according to the invention are preferably capable of at least one of the following reactions:

a) oxidation of unsubstituted or substituted N-, O- or S-heterocyclic mono-, bi- or polynuclear aromatic compounds;
b) oxidation of unsubstituted or substituted mono- or polynuclear aromatics;
c) oxidation of straight-chain or branched alkanes and alkenes; and
d) oxidation of unsubstituted or substituted cycloalkanes and cycloalkenes.

Preferred monooxygenase mutants have at least one functional mutation, in particular amino acid substitution, in at least one of the sequence regions 73-82, 86-88 and 172-224. Thus, for example, Phe87 can be replaced by an amino acid having an aliphatic side chain, such as Ala, Val, Leu, in particular Val; Leu188 can be replaced by an amino acid having an amide side chain, such as Asn or, in particular, Gln; and Ala74 can be

replaced by another amino acid having an aliphatic side chain, such as Val and, in particular, Gly.

Particularly preferred monooxygenase mutants of this type are those which have at least one of the following mono- or polyamino acid substitutions:

1) Phe87Val;
2) Phe87Val, Leu188Gln; or
3) Phe87Val, Leu188Gln, Ala74Gly;

and functional equivalents thereof. The number indicates the position of the mutation; the original amino acid is indicated before the number and the newly introduced amino acid after the number.

In this context, "functional equivalents" or analogs of the mutants which are disclosed specifically are mutants differing therefrom which furthermore have the desired substrate specificity with respect to at least one of the oxidation reactions a) to d) described above, i.e., for example, for heterocyclic aromatics and which hydroxylate, for example, indole, or furthermore exhibit the desired "modified substrate profile" with respect to the wild-type enzyme.

"Functional equivalents" are also to be understood as meaning in accordance with the invention mutants which exhibit, in at least one of the abovementioned sequence positions, an amino acid substitution other than the one mentioned specifically, but still lead to a mutant which, like the mutant which has been mentioned specifically, show a "modified substrate profile" with respect to the wild-type enzyme and catalyze at least one of the abovementioned oxidation reactions. Functional equivalence exists in particular also in the case where the modifications in the substrate profile correspond qualitatively, i.e. where, for example, the same substrates are converted, but at different rates.

"Functional equivalents" naturally also encompass P450 monooxygenase mutants which, like the P450 BM3 mutants which have been mentioned specifically, can be obtained by mutating P450 enzymes from other organisms. For example, regions of homologous sequence regions can be identified by sequence comparison. Following the principles of what has been set out specifically in the invention, the modern methods of molecular modeling then

allow equivalent mutations to be carried out which affect the reaction pattern.

"Functional equivalents" also encompass the mutants which can be obtained by one or more additional amino acid additions, substitutions, deletions and/or inversions, it being possible for the abovementioned additional modifications to occur in any sequence position as long as they give rise to a mutant with a modified substrate profile in the above sense.

Substrates of group a) which can be oxidized according to the invention are unsubstituted or substituted heterocyclic mono-, bi- or polynuclear aromatic compounds; in particular oxidizable or hydroxylatable N-, O- or S-heterocyclic mono-, bi- or polynuclear aromatic compounds. They include preferably two or three, in particular two, 4- to 7-membered, in particular 6- or 5-membered, fused rings, where at least one, preferably all, rings have aromatic character and where at least one of the aromatic rings carries one to three, preferably one, N-, O- or S-heteroatom in the ring. The total ring structure may contain one or two further identical or different heteroatoms. The aromatic compounds may furthermore carry 1 to 5 substituents at the ring carbon or heteroatoms. Examples of suitable substituents are C_1- to C_4-alkyl, such as methyl, ethyl, n- or isopropyl, n-, iso- or t-butyl, or C_2- to C_4-alkenyl, such as ethenyl, 1-propenyl, 2-propenyl, 1-butenyl, 2-butenyl or 3-butenyl, hydroxyl and halogen, such as F, Cl and Br. The alkyl or alkenyl substituents mentioned may also have a keto or aldehyde group; examples being propan-2-on-3-yl, butan-2-on-4-yl, 3-buten-2-on-4-yl. Non-limiting examples of suitable heterocyclic substrates are, in particular, binuclear heterocycles, such as indole, N-methyl-indole, and the substituted analogs thereof which carry one to three of the above-defined substituents on carbon atoms, for example 5-chloro- or 5-bromoindole; and also quinoline and quinoline derivatives, for example 8-methylquinoline, 6-methyl-quinoline and quinaldine; and benzothiophene, and the substituted analogs thereof which carry one to three of the above-defined substituents on carbon atoms. Moreover, trinuclear hetero-aromatics, such as acridine and the substituted analogs thereof which carry one to three of the above-defined substituents on carbon atoms, may be mentioned.

Substrates of group b) which are oxidizable according to the invention are unsubstituted or substituted mono- or polynuclear, in particular mono- or binuclear, aromatics, such as benzene and naphthalene. The aromatic compounds may be unsubstituted or mono- or polysubstituted and, for example, carry 1 to 5 substituents on

the ring carbon atoms. Examples of suitable substituents are C_1- to C_4-alkyl, such as methyl, ethyl, n- or isopropyl or n-, iso- or t-butyl, or C_2- to C_4-alkenyl, such as ethenyl, 1-propenyl, 2-propenyl, 1-butenyl, 2-butenyl or 3-butenyl, hydroxyl and halogen, such as F, Cl and Br. The alkyl or alkenyl substituents mentioned may also have a keto or aldehyde group; Examples being propan-2-on-3-yl, butan-2-on-4-yl, 3-buten-2-on-4-yl. The aromatic may be fused with a four- to seven-membered non-aromatic ring. The non-aromatic ring may have one or two C=C double bonds, be mono- or polysubstituted by the abovementioned substituents and may carry one or two hetero ring atoms. Examples of particularly suitable aromatics are mononuclear aromatics, such as cumene, and binuclear substrates, such as indene and naphthalene, and substituted analogs thereof which carry one to three of the above-defined substituents on carbon atoms.

Substrates of group c) which can be oxidized according to the invention are straight-chain or branched alkanes or alkenes having 4 to 15, preferably 6 to 12, carbon atoms. Examples which may be mentioned are n-butane, n-pentane, n-hexane, n-heptane, n-octane, n-nonane, n-decane, n-undecane and n-dodecane, and the analogs of these compounds which are branched once or more than once, for example analogous compounds having 1 to 3 methyl side groups; or mono- or polyunsaturated, for example mono-unsaturated, analogs of the abovementioned alkanes.

Substrates of group d) which can be oxidized according to the invention are unsubstituted or substituted cycloalkanes and cycloalkenes having 4 to 8 ring carbon atoms. Examples of these are cyclopentane, cyclopentene, cyclohexane, cyclohexene, cycloheptane and cycloheptene. The ring structure may carry one or more, for example 1 to 5, substituents according to the above definition for compounds of groups a) and b). Nonlimiting examples are ionones, such as α-, β- and γ-ionone, and the corresponding methyl ionones and iso-methyl ionones. Particular preference is given to α- and β-ionone.

The invention also relates to nucleic acid sequences coding for one of the monooxygenases according to the invention. Preferred nucleic acid sequences are derived from SEQ ID NO:1, which have at least one nucleotide substitution which leads to one of the functional amino acid mutations described above. The invention moreover relates to functional analogs of the nucleic acids obtained by addition, substitution, insertion and/or deletion of individual or multiple nucleotides, which furthermore code for a

monooxygenase having the desired substrate specificity, for example having indole-oxidizing activity.

The invention also encompasses those nucleic acid sequences which comprise so-called silent mutations or which are modified in comparison with a specifically mentioned sequence in accordance with the codon usage of a specific origin or host organism, and naturally occurring variants of such nucleic acid sequences. The invention also encompasses modifications of the nucleic acid sequences obtained by degeneration of the genetic code (i.e. without any changes in the corresponding amino acid sequence) or conservative nucleotide substitution (i.e. the corresponding amino acid is replaced by another amino acid of the same charge, size, polarity and/or solubility), and sequences modified by nucleotide addition, insertion, inversion or deletion, which sequences encode a monooxygenase according to the invention having a "modified substrate profile", and the corresponding complementary sequences.

The invention furthermore relates to expression constructs comprising a nucleic acid sequence encoding a mutant according to the invention under the genetic control of regulatory nucleic acid sequences; and vectors comprising at least one of these expression constructs.

Preferably, the constructs according to the invention encompass a promoter 5'-upstream of the encoding sequence in question and a terminator sequence 3'-downstream, and, optionally, further customary regulatory elements, and, in each case operatively linked with the encoding sequence. Operative linkage is to be understood as meaning the sequential arrangement of promoter, encoding sequence, terminator and, if appropriate, other regulatory elements in such a manner that each of the regulatory elements can fulfill its intended function on expression of the encoding sequence. Examples of operatively linkable sequences are targeting sequences, or else translation enhancers, enhancers, polyadenylation signals and the like. Further regulatory elements encompass selectable markers, amplification signals, replication origins and the like.

In addition to the artificial regulatory sequences, the natural regulatory sequence can still be present upstream of the actual structural gene. If desired, this natural regulation may be switched off by genetic modification, and the expression of the genes may be enhanced or lowered. However, the gene construct may also be simpler in construction, i.e. no additional regulatory signals are inserted upstream of the structural gene and the

natural promoter with its regulation is not removed. Instead, the natural regulatory sequence is mutated in such a way that regulation no longer takes place and the gene expression is increased or reduced. One or more copies of the nucleic acid sequences may be present in the gene construct.

Examples of suitable promoters are: cos, tac, trp, tet, trp-tet, lpp, lac, lpp-lac, lacIq, T7, T5, T3, gal, trc, ara, SP6, l-PR or l-PL promoter, which are advantageously employed in Gram-negative bacteria; and Gram-positive promoters amy and SPO2, the yeast promoters ADC1, MFa, Ac, P-60, CYC1, GAPDH or the plant promoters CaMV/35S, SSU, OCS, lib4, usp, STLS1, B33, nos or the ubiquitin or phaseolin promoter. Particular preference is given to using inducible promoters, for example light- and in particular temperature-inducible promoters, such as the P_rP_l promoter.

In principle, all natural promoters with their regulatory sequences can be used. In addition, synthetic promoters may also be used in an advantageous fashion.

The abovementioned regulatory sequences are intended to allow the targeted expression of the nucleic acid sequences and of protein expression. Depending on the host organism, this may mean, for example, that the gene is expressed or overexpressed only after induction has taken place, or that it is expressed and/or overexpressed immediately.

The regulatory sequences or factors can preferably have a positive effect on expression and in this manner increase or reduce the latter. Thus, an enhancement of the regulatory elements may advantageously take place at the transcriptional level by using strong transcription signals such as promoters and/or "enhancers". In addition, translation may also be enhanced by improving, for example, mRNA stability.

An expression cassette is generated by fusing a suitable promoter with a suitable monooxygenase nucleotide sequence and a terminator signal or polyadenylation signal. To this end, customary recombination and cloning techniques are used as they are described, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) and in T.J. Silhavy, M.L. Berman and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and in Ausubel, F.M. et al., Current Protocols in

Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987).

For expression in a suitable host organism, the recombinant nucleic acid construct or gene construct is advantageously inserted into a host-specific vector which allows optimal gene expression in the host. Vectors are well known to the skilled worker and can be found, for example, in "Cloning Vectors" (Pouwels P.H. et al., Ed., Elsevier, Amsterdam-New York-Oxford, 1985). Vectors are to be understood as meaning not only plasmids, but all other vectors known to the skilled worker such as, for example, phages, viruses, such as SV40, CMV, baculovirus and adenovirus, transposons, IS elements, phasmids, cosmids, and linear or circular DNA. These vectors can be replicated autonomously in the host organism or chromosomally.

The vectors according to the invention allow the generation of recombinant microorganisms which are transformed, for example, with at least one vector according to the invention and which can be employed for producing the mutants. The above-described recombinant constructs according to the invention are advantageously introduced into a suitable host system and expressed. It is preferred to use usual cloning and transfection methods known to the skilled worker in order to bring about expression of the abovementioned nucleic acids in the expression system in question. Suitable systems are described, for example, in current protocols in molecular biology, F. Ausubel et al., Ed., Wiley Interscience, New York 1997.

Suitable host organisms are, in principle, all organisms which allow expression of the nucleic acids according to the invention, their allelic variants, and their functional equivalents or derivatives. Host organisms are to be understood as meaning, for example, bacteria, fungi, yeasts or plant or animal cells. Preferred organisms are bacteria such as those of the genera Escherichia, such as, for example, Escherichia coli, Streptomyces, Bacillus or Pseudomonas, eukaryotic microorganisms such as Saccharomyces cerevisiae, Aspergillus, and higher eukaryotic cells from animals or plants, for example Sf9 or CHO cells.

If desired, expression of the gene product may also be brought about in transgenic organisms such as transgenic animals such as, in particular, mice, sheep, or transgenic plants. The transgenic organisms may also be knock-out animals or plants in which the

corresponding endogenous gene has been eliminated, such as, for example, by mutation or partial or complete deletion.

Successfully transformed organisms can be selected by marker genes which are likewise contained in the vector or in the expression cassette. Examples of such marker genes are genes for resistance to antibiotics and for enzymes which catalyze a color reaction, which causes staining of the transformed cell. These transformed cells can then be selected using automatic cell selection. Microorganisms which have been transformed successfully with a vector and which carry an appropriate gene for resistance to antibiotics (for example G418 or hygromycin) can be selected by using appropriate antibiotics-containing media or substrates. Marker proteins which are presented on the cell surface can be used for selection by affinity chromatography.

The combination of the host organisms and the vectors appropriate for the organisms, such as plasmids, viruses or phages, such as, for example, plasmids with the RNA polymerase/promoter system, phages λ, μ or other temperate phages or transposons and/or other advantageous regulatory sequences forms an expression system. The term "expression system" means, for example, a combination of mammalian cells such as CHO cells, and vectors, such as pcDNA3neo vector, which are suitable for mammalian cells.

As described above, the gene product can also be expressed advantageously in transgenic animals, for example mice, sheep, or transgenic plants. It is likewise possible to program cell-free translation systems with the RNA derived from the nucleic acid.

The invention furthermore provides a process for preparing a monooxygenase according to the invention, which comprises cultivating a monooxygenase-producing microorganism, if appropriate inducing the expression of the monooxygenase, and isolating the monooxygenase from the culture. If desired, the monooxygenase according to the invention can thus also be produced on an industrial scale.

The microorganism can be cultivated and fermented by known methods. Bacteria, for example, can be grown in a TB or LB medium and at 20-40°C and a pH of 6-9. Suitable cultivation conditions are described in detail in T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989), for example.

If the monooxygenase is not secreted into the culture medium, the cells are then lyzed and the monooxygenase is obtained from the lysate using known methods for the isolation of proteins. The cells can be lyzed alternatively by high-frequency ultrasound, by high pressure, for example in a French pressure cell, by osmolysis, by the action of detergents, lytic enzymes or organic solvents, by homogenization or by a combination of a plurality of the processes mentioned. Purification of the monooxygenase can be achieved by known chromatographic processes, such as molecular sieve chromatography (gel filtration), such as Q-Sepharose chromatography, ion-exchange chromatography and hydrophobic chromatography, and by other customary processes, such as ultrafiltration, crystallization, salting out, dialysis and native gel electrophoresis. Suitable processes are described, for example, in Cooper, F.G., Biochemische Arbeitsmethoden [Biochemical Procedures], Verlag Walter de Gruyter, Berlin, New York or in Scopes, R., Protein Purification, Springer Verlag, New York, Heidelberg, Berlin.

To isolate the recombinant protein, it is particularly advantageous to use vector systems or oligonucleotides which extend the cDNA by certain nucleotide sequences and thus code for modified polypeptides or fusion proteins which serve to simplify purification. Suitable modifications of this type are, for example, so-called "tags" which act as anchors, such as, for example, the modification known as hexa-histidine anchor, or epitopes which can be recognized as antigens by antibodies (described, for example, in Harlow, E. and Lane, D., 1988, Antibodies: A Laboratory Manual. Cold Spring Harbor (N.Y.) Press). These anchors can be used to attach the proteins to a solid support such as, for example, a polymer matrix, which can, for example, be packed into a chromatography column, or to a microtiter plate or to another support.

These anchors can also at the same time be used to recognize the proteins. It is also possible to use for recognition of the proteins conventional markers such as fluorescent dyes, enzyme markers which form a detectable reaction product after reaction with a substrate, or radioactive markers, alone or in combination with the anchors for derivatizing the proteins.

The invention moreover relates to a process for the microbiological oxidation of organic compounds, for example N-heterocyclic mono-, bi- or polynuclear aromatic compounds according to the above definition, which comprises

a1) culturing a recombinant microorganism according to the above definition in a culture medium, in the presence of an exogenous (added) substrate or an intermediately formed substrate, which substrate is oxidizable by the monooxygenase according to the invention, preferably in the presence of oxygen (i.e. aerobically); or

a2) incubating a substrate-containing reaction medium with an enzyme according to the invention, preferably in the presence of oxygen and an electron donor; and

b) isolating the oxidation product formed or a secondary product thereof from the medium.

The oxygen required for the reaction either passes from the atmosphere into the reaction medium or, if required, can be added in a manner known per se.

The oxidizable substrate is preferably selected from

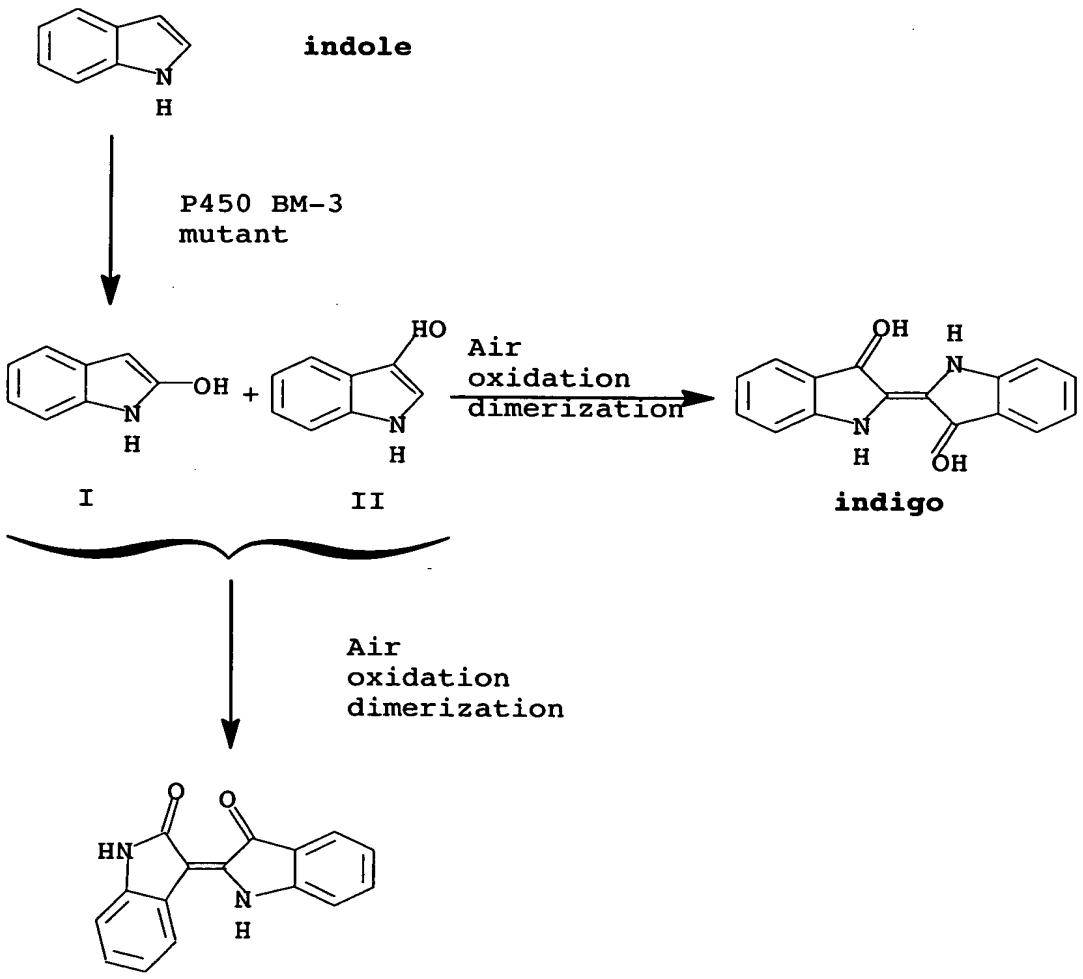
a) unsubstituted or substituted N-heterocyclic mono-, bi- or polynuclear aromatic compounds;
b) unsubstituted or substituted mono- or polynuclear aromatics;
c) straight-chain or branched alkanes and alkenes;
d) unsubstituted or substituted cycloalkanes and cycloalkenes.

A preferred process variant is directed to the formation of indigo/indirubin and is characterized by the fact that the substrate is indole formed as an intermediate in the culture and that the indigo and/or indirubin formed in the culture medium is isolated by oxidation of hydroxyindole intermediates.

If the oxidation according to the invention is carried out using a recombinant microorganism, the culturing of the microorganisms is preferably first carried out in the presence of oxygen and in a complex medium, such as, for example, TB or LB medium at a culturing temperature of approximately 20 to 40°C and a pH of approximately 6 to 9, until an adequate cell density is reached. The addition of exogenous indole is usually not necessary, as this is intermediately formed by the microorganism. However, when using other substrates, addition of exogenous substrate may be required. In order to be able to control the oxidation reaction better, the use of an inducible, in particular temperature-inducible, promoter is preferred. The temperature is in this case increased to the necessary induction temperature, e.g. 42°C in the case of the P_rP_l promoter, this is maintained for a sufficient period of time, e.g. 1 to 10 or 5 to 6 hours, for the expression of the monooxygenase activity and the temperature is then reduced again to a value of approximately 30 to 40°C. The culturing is

then continued in the presence of oxygen for 12 hours to 3 days. The pH can, in particular in the case of indole oxidation, be increased by addition of NaOH, e.g. to 9 to 10, whereby the indigo formation or indirubin formation is additionally promoted by atmospheric oxidation of the enzymatically formed oxidation products 2- and 3-hydroxyindole.

The indigo/indirubin formation according to the invention is illustrated by the reaction scheme below:



I: 2-hydroxyindole (oxindole)
II: 3-hydroxyindole (indoxyl)

However, if the oxidation according to the invention is carried out using purified or enriched enzyme mutants, the enzyme according to the invention is dissolved in an exogenous

substrate-containing, for example indole-containing medium (approximately 0.01 to 10 mM, or 0.05 to 5 mM), and the reaction is carried out, preferably in the presence of oxygen, at a temperature of approximately 10 to 50°C, such as, for example, 30 to 40°C, and a pH of approximately 6 to 9 (such as established, for example, using 100 to 200 mM phosphate or tris buffer), and in the presence of a reductant, the substrate-containing medium moreover containing, relative to the substrate to be oxidized, an approximately 1- to 100-fold or 10- to 100-fold molar excess of reduction equivalents. The preferred reductant is NADPH. If required, the reducing agent can be added in portions.

In a similar manner, the oxidizable substrates which are preferably used are: n-hexane, n-octane, n-decane, n-dodecane, cumene, 1-methylindole, 5-Cl- or Br-indole, indene, benzothiophene, α-, β- and γ-ionone, acridine, naphthalene, 6-methyl- or 8-methylquinoline, quinoline and quinaldine.

The enzymatic oxidation reaction according to the invention can be carried out, for example, under the following conditions:

Substrate concentration:	from 0.01 to 20 mM
Enzyme concentration:	from 0.1 to 10 mg/ml
Reaction temperature:	from 10 to 50°C
pH:	from 6 to 8
Buffer:	from 0.05 to 0.2 M potassium phosphate, or Tris/HCl
Electron donor:	is preferably added in portions (initial concentration about 0.1 to 2 mg/ml)

The mixture can briefly (from 1 to 5 minutes) be preincubated (at about 20-40°C) before the reaction is initiated, for example by adding the electron donors (e.g. NADPH). The reaction is carried out aerobically, if appropriate with additional introduction of oxygen.

In the substrate oxidation process according to the invention, oxygen which is present in or added to the reaction medium is cleaved reductively by the enzyme. The required reduction

equivalents are provided by the added reducing agent (electron donor).

The oxidation product formed can then be separated off from the medium and purified in a conventional manner, such as, for example, by extraction or chromatography.

Further subjects of the invention relate to bioreactors, comprising an enzyme according to the invention or a recombinant microorganism according to the invention in immobilized form.

A last subject of the invention relates to the use of a cytochrome P450 monooxygenase according to the invention or of a vector or microorganism according to the invention for the microbiological oxidation of a substrate from one of the groups a) to d), in particular of N-heterocyclic mono-, bi- or polynuclear aromatic compounds, and preferably for the formation of indigo and/or indirubin.

The present invention is now described in greater detail with reference to the following examples.

Example 1:

Randomization of specific codons of P450 BM-3

The experiments were carried out essentially as described in (19). Three positions (Phe87, Leu188 and Ala74) were randomized with the aid of site-specific mutagenesis using the Stratagene QuikChange kit (La Jolla, CA, USA). The following PCR primers were used for the individual positions:

Phe87: 5'-gcaggagacgggttgnnnacaagctggacg-3' (SEQ ID NO:3),
5'-cgtccagcttgtnnncaacccgtctcctgc-3', (SEQ ID NO:4)
Leu188: 5'-gaagcaatgaacaagnnncagcgagcaaatccag-3' (SEQ ID NO:5),
5'-ctggatttgctcgctgnnncttgttcattgcttc-3' (SEQ ID NO:6);
Ala74: 5'-gctttgataaaaacttaaagtcaannncttaaatttgtacg-3' (SEQ ID: NO:7),
5'-cgtacaaatttaagnnnttgacttaagtttttatcaaagc-3' (SEQ ID NO:8)

The conditions for the PCR were identical for all three positions. In particular, 17.5 pmol of one of each primer, 20 pmol of template plasmid DNA, 3 U of the Pfu polymerase and 3.25 nmol of each dNTP were used per 50 µl reaction volume. The PCR reaction was started at 94°C/1 min and the following temperature cycle was then carried out 20 times: 94°C, 1 min;

46°C, 2.5 min; 72°C, 17 min. After 20 cycles, the reaction was continued at 72°C for 15 min. After the PCR, the template DNA was digested at 37°C for 3 h using 20 U of DpnI. *E. coli* DH5α was then transformed. The transformed *E. coli* DH5α cells were plated out onto LB agar plates which contained 150 µg/ml of ampicillin. Incubation was then carried out at 37°C for 18 h.

Example 2:

Expression and purification of the P450 BM-3 and its mutants and production of a blue pigment

The P450 BM-3 gene and the mutants thereof were expressed under the control of the strong, temperature-inducible P_RP_L promoter of the plasmid pCYTEXP1 in *E. coli* DH5α as already described (20). Colonies were picked up using sterile toothpicks and transferred to microtiter plates having 96 hollows, comprising 200 µl of TB medium and 100 µg/ml of ampicillin per hollow. Incubation was then carried out at 37°C overnight. 40 µl of the cell culture of one of each hollow were then transferred to a culture tube which contained 2 ml of TB medium with 100 µg/ml of ampicillin. Culturing was then carried out at 37°C for 2 h. The temperature was then increased to 42°C for 6 h for induction. Culturing was then continued at 37°C overnight, a blue pigment being produced.

The preparative production of enzyme or blue pigment was carried out starting from a 300 ml cell culture (OD_{578nm}= 0.8 to 1.0). For the isolation of the enzyme, the cells were centrifuged off at 4000 rpm for 10 min and resuspended in 0.1 M K_xPO_4 buffer, pH 7.4. The ice-cooled cells were carefully disrupted with the aid of a Branson sonifer W25 (Dietzenbach, Germany) at an energy output of 80 W by 2 min sonification three times. The suspensions were centrifuged at 32570 x g for 20 min. The crude extract was employed for the activity determination or for the enzyme purification. The enzyme purification was carried out as already described in (21), to which reference is expressly made hereby. The concentration of purified enzyme was determined by means of the extinction difference at 450 and 490 nm, as already described in (11), using an extinction coefficient ε of 91 mM^{-1} cm^{-1}.

Example 3:

Isolation of mutants which produce large amounts of blue pigment

100 colonies in each case were isolated from the mutants of one of each position, which were produced by randomized mutagenesis of the codon of the corresponding position. These colonies were cultured in culture tubes for the production of blue pigment.

After washing the cells with water and a number of slow centrifugation steps (500 rpm), the blue pigment was extracted using dimethyl sulfoxide (DMSO). The solubility of the blue pigment was greatest in DMSO. The absorption of the extract was determined at 677 nm. That mutant which produced the largest amount of blue pigment, especially mutants from a specific position, was used for DNA sequencing (ABI DNA sequencing kit; ABI Prism™ 377 DNA sequencer) and moreover as a template for site-specific randomized mutagenesis.

Example 4:

Activity test for the indole hydroxylation

The indole hydroxylation activity was tested in a solution which contained 8 µl of a 10-500 mM indole solution in DMSO, 850 µl of tris/HCl buffer (0.1 M, pH 8.2) and 0.6 nmol of P450 BM-3 wild type or mutant in a final volume of 1 ml. The mixture was preincubated for 9 min before the reaction was started by addition of 50 µl of an aqueous 1 mM solution of NADPH. The reaction was stopped after 20 sec by addition of 60 µl of 1.2 M KOH. Within 5 to 30 sec (under aerobic conditions), the enzyme products were converted completely into indigo [$\Delta^{2,2'}$-biindoline]-3,3'-dione) and indirubin ([$\Delta^{2,3'}$-biindoline]-2',3-dione). The indigo production was determined by means of its absorption at 670 nm. A calibration curve using pure indigo showed an extinction coefficient of 3.9 mM^{-1} cm^{-1} at this wavelength. A linear curve was obtained for indigo production in a reaction time of 40 sec using 0.6 nmol of wild type or P450 BM-3 mutant and 0.05 to 5.0 mM of indole. Indirubin shows a very weak absorption at 670 nm and the amount of indirubin formed was very much smaller than the amount of indigo formed. The formation of indirubin was neglected in the determination of the kinetic parameters. The NADPH consumption was determined at 340 nm and calculated as described (17) using an extinction coefficient of 6.2 mM^{-1} cm^{-1}.

Example 5:

Purification of indigo and indirubin

After washing the cells with water and repeated centrifugation at 500 g, the blue pellet formed was extracted using tetrahydrofuran (THF). The extract was evaporated almost to dryness and the red pigment was extracted a number of times with 50 ml of absolute ethanol. The residual blue solid was dissolved in THF and analyzed by thin-layer chromatography (TLC). The ethanol solution

was evaporated and purified by silica gel chromatography (TLC 60, Merck, Darmstadt, Germany; 2 cm x 30 cm) before it was washed with THF and petroleum ether in a ratio of 1:2. The red solution obtained was evaporated and its purity was determined by TLC. The absorption spectra of the blue and of the red pigment were determined in a range from 400 to 800 nm with the aid of an Ultraspec 3000 spectrophotometer (Pharmacia, Uppsala, Sweden). The blue and the red color were moreover analyzed by mass spectrometry and ^{1}H-NMR spectroscopy.

Experimental results

1. Increasing the productivity for blue pigment by P450 BM-3 mutagenesis

Native P450 BM-3 does not have the ability to produce the blue indigo-containing pigment, or the precursor substances 2- or 3-hydroxyindole. In order to be able to prepare a sufficient amount of blue pigment, P450 BM-3 was subjected to evolution in a controlled manner. All mutants which produced the blue pigment were sequenced. It was found that at least one of the following three positions were mutated: Phe87, Leu188 and Ala74. It was therefore assumed that these three positions play a crucial role for the activity of P450 BM-3 in the production of blue pigment. From the structure of the heme domain of cytochrome P450 BM-3, complexed with palmitoleic acid, it is seen that Phe87 prevents the substrate from coming closer to the heme group (14). The mutant Phe87Val shows a high regio- and stereoselectivity in the epoxidation of (14S, 15R)-arachidonic acid (13) and the mutant Phe87Ala shifts the hydroxylation position of ω-1, ω-2 and ω-3 to ω (22). The position 87 was therefore selected as first for the site-specific randomized mutagenesis with the aid of PCR. In tube cultures, 7 colonies were obtained which produced a small amount of blue pigment after induction. The colony which produced the largest amount of the blue pigment was selected for the DNA sequencing. The sequence data showed substitution of Phe87 by Val. The mutant Phe87Val was then used as a template for the second round of site-specific randomized mutagenesis on position Leu188. The structure of the heme domain, complexed with palmitoleic acid, shows that the repositioning of the F and G helices brings the residue Leu188 into direct contact with the substrate (14). This position can therefore play an important role in substrate binding or orientation. After the second screening passage, 31 colonies were observed which produced the blue pigment. The mutant which produced the largest amount of pigment contained the substitutions Phe87Val and Leu188Gln. This mutant was then mutated in position Ala74 in the third passage of

site-specific randomized mutagenesis. In this case the triple mutant F87L188A74 (Phe87Val, Leu188Gln and Ala74Gly) was obtained, which produced several mg of blue pigment in a 2-liter flask, containing 300 ml of TB medium. This amount was sufficient for the isolation and characterization of the blue pigment.

2. Isolation and identification of the blue pigment

After washing the cells, the residual blue pellet was extracted with THF and analyzed by TLC. The blue pigment was separated into a rapidly migrating blue component and into a more slowly migrating red component. Both components showed exactly the same mobility parameters as the components of a commercial indigo sample.

After the purification, the absorption spectra of both components were determined in DMSO. The blue component showed the same spectrum as a commercial indigo sample. The purified blue and red components were each analyzed by mass spectrometry. The mass spectra of both pigments showed a strong molecular ion peak at m/e = 262 and two fragment peaks at m/e = 234 and 205 (relative intensity in each case 10%). This pattern is typical of indigoid compounds. The elementary composition of these ions was determined by high-resolution mass spectrometry as $C_{16}H_{10}N_2O_2$, $C_{15}H_{10}N_2O$ and $C_{14}H_9N_2$. This is also characteristic of structures of the indigo type. The blue pigment was thus identified as indigo and the red pigment as indirubin. For the confirmation of the structure, 500 MHz ^{1}H-NMR spectra of both pigments were carried out in DMSO-D_6 solution. The results agreed with the literature data (23).

3. Production of indigo using isolated enzymes

It is known that indigo is accessible from indole by microbial transformation (24-26). None of these microbial systems, however, contained a P450 monooxygenase. According to the invention, the catalytic activity of the pure enzyme for indole was first determined. The mutant F87L188A74 was mixed with indole. No color reaction could be observed. Only after addition of NADPH to the reaction mixture was the blue pigment formed after approximately 20 min. By adjustment of the pH of the reaction mixture to a value of approximately 11, 30 sec after addition of NADPH, the blue coloration was visible within a few seconds. Control experiments using native P450 BM-3 were always negative, even using increased concentrations of enzyme, indole and NADPH. The blue pigment was extracted using ethyl acetate and analyzed by TLC. The blue pigment again separated into a more rapidly running

blue component and into a slower running red component. The Rf values and the absorption spectra were identical to those values of the extracts from the fermentation broth. The F87L188A74 mutant of P450 BM-3 is thus an indole hydroxylase.

Two routes have previously been described for the enzymatic transformation of indole to indigo. One route is catalyzed by a dioxygenase, the other by a styrene monooxygenase (24, 25). The NADPH stoichiometry is in both cases 2. It was therefore assumed that in contrast to the dioxygenases the mutant F87L188A74 according to the invention hydroxylates indole in only one position to form oxindole (2-hydroxyindole) or indoxyl (3-hydroxyindole).

4. Kinetic parameters of indole hydroxylation

Pure samples of the wild-type enzyme P450 BM-3 and of the mutants Leu188Gln, Phe87Val, F87L188 and F87L188A74 were used for the determination of the kinetic parameters of indole hydroxylation. The results are summarized in Table 1 below.

Table 1: Kinetic parameters of the P450 BM-3 mutants for indole hydroxylation

Mutants	$K_{cat}(S^{-1})$	K_m (mM)	K_{cat}/K_m ($M^{-1}s^{-1}$)
WT	-[a]	-	-
Leu188Gln	n.d.[b]	n.d.	n.d.
Phe87Val	2.03 (0.14)	17.0 (1.0)	119
F87L188	2.28 (0.16)	4.2 (0.4)	543
F87L188A74	2.73 (0.16)	2.0 (0.2)	1365

a) no activity was observed;
b) not determined (activity was too low to be measured)

Even with an excess of purified enzyme and high indole concentration, the wild-type enzyme is not able to oxidize indole. The mutant Leu188Gln shows a low activity. The mutant Phe87Val shows a catalytic activity of 119 $M^{-1}s^{-1}$ for indole hydroxylation. The catalytic efficiency of the double mutant F87L188 (Phe87Val,Leu188Gln) increased to 543 $M^{-1}s^{-1}$ and was increased to 1365 $M^{-1}s^{-1}$ by introduction of the further substitution Ala74Gly. The K_{cat} values increased from Phe87Val to the triple mutant by a total of 35%, while the K_m values decreased approximately by seven-fold. This indicates that Ala74Gly and Leu188Gln are mainly involved in substrate binding.

For the triple mutant F87L188A74, the indole turnover rate (K_{cat}=2.73 s^{-1}) was more than ten times higher than for most P450 enzymes (18).

Example 6

Hydroxylation of n-octane using modified cytochrome P450 monooxygenase

The reactions were carried out using a P450 BM-3 monooxygenase mutant comprising the following mutations: Phe87Val Leu188Gln Ala74Gly

The chosen substrate was n-octane. For the hydroxylation of n-octane, the following aerobic reaction mixture was used:

P450 BM-3 mutant:	17.5 mg (lyophilisate)
Reaction buffer:	9.1 ml (potassium phosphate buffer 50 mM, pH 7.5)
Substrate:	50 µl of a 60 mM solution (in acetone)
Temperature:	25°C

The enzyme lyophilisate was dissolved in 500 µl of reaction buffer and initially incubated at room temperature with substrate and reaction buffer for 5 minutes. 300 µl NADPH solution (5 mg/ml) were then added. Addition of NADPH was repeated two more times. The progress of the reaction was monitored by measuring the absorption at 340 nm, which allows the NADPH decrease to be observed. NADPH is added in aliquots of 300 µl, since too high a concentration of NADPH in the reaction solution would result in inactivation of the enzyme. To isolate the products, the reaction solution was then extracted three times with 5 ml of diethyl ether. The combined organic phases were dried over $MgSO_4$ and concentrated. The products were then characterized by TLC, GC/MS and NMR.

The GC/MS analysis of the reaction mixture gave the following result:

Compound	Rt[min][1)]	Conversion [%]
4-octanol	13.51	37
3-octanol	14.08	47
2-octanol	14.26	16

1) Temperature program: 40°C 1 min isothermic / 3°C/min 95°C /10°C/min 275°C; apparatus: Finnigan MAT 95; GC: HP 5890 Series II

Split Injector; Column: HP-5MS (methylsiloxane) 30m x 0.25mm; Carrier gas: 0.065 ml of He/min

No starting material was found.

Example 7:

Hydroxylation of aromatics, heteroaromatics and trimethylcyclohexenyl compounds

a) Example 6 was repeated, but using, instead of n-octane, the substrate naphthalene. The products that were identified were 1-naphthol and cis-1,2-dihydroxy-1,2-dihydronaphthalene. 88% of the naphthalene starting material had been converted.

Analytic methods for reactions with naphthalene

GC:
Apparatus: Carlo Erba Strumentazion Typ HRGC 4160 on Column Injector; Column: DB5 30m x 0.2 mm; Material: 5% diphenyl-95% dimethylpolysiloxane; Carrier gas: 0.5 bar H_2; Temperature program: 40°C 1 min isothermic / 10°C/min to 300°C
Rt(1-naphthol) = 16.68

NMR:
1-Naphthol and cis-1,2-dihydroxy-1,2-dihydro-naphthalene were identified in the ^1H NMR.

b) Example 6 was repeated but using, instead of n-octane, the substrate 8-methylquinoline. 5-Hydroxy-8-methylquinoline was identified as main product, in addition to other derivatives (product ratio 5:1). 35% of the starting material used had been converted.

c) Example 6 was repeated but using, instead of n-octane, the substrate α-ionone. 3-Hydroxy-α-ionone was identified as main product, in addition to other derivatives (product ratio: 76:24). 60% of the starting material used had been converted.

d) Example 6 was repeated, but using, instead of n-octane, the substrate cumene (isopropylbenzene). Five monohydroxy products and one dihydroxy product were identified. 70% of the starting material used had been converted.

REFERENCES

1. Yano, T., Oue, S., and Kagamiyama, H. (1998) Proc. Natl. Acad Sci. USA 95, 5511-5515.

2. Zhang, J.-H., Dawes, G., and Stemmer, W. P. C. (1997) Proc. Natl. Acad Sci. USA 94, 4504-4509.

3. Wan, L., Twitchett, M. B., Eltis, L. D., Mauk, A. G., and Smith, M. (1998) Proc. Natl. Acad Sci USA 95, 12825-12831.

4. Cronin, C. N. (1998) J. Biol. Chem 273, 24465-24469.

5. Wilks, H. M., Hart, K. W., Feeney, R., Dunn, C. R., Muirhead, H., Chia, W. N., Barstow, D. A., Atkinson, T., Clarke, A. R., Holbrook, I J. (1988) Science 242, 1541-1544.

6. Hedstrom, L., Szilagyi, L., Rutter, W. J. (1992) Science 255, 1249-1253.

7. Tucker, C. L., Hurley, J. H., Miller, T. R., and Hurley, I B. (1998) Proc. Natl. Acad Sci. USA 95, 5993-5997.

8. Quemeneur, E., Moutiez, J.-B. C., and Menez, A. (1998) Nature (London) 391, 301-303.

9. Marsden, A- F. A., Wilkinson, B., Cortes, J., Dunster, N. J., Staunton, I Leadlay, P. F. (1998) Science 279, 199-201.

10. Chen, R., Greer, A., and Dean, A. M. (1998) Proc. Natl. Acad Sci. US4 95, 11666-11670.

11. Boddupalli, S. S., Estabrook, R. W. and Peterson, J. A. (1990) J Biol. Chem 265, 4233-4239.

12. Capdevila, J. H., Wie, S., Helvig, C., Falck, J. R., Belosludtsev, Y.,Truan, G., Graham-Lorence, S. E., and Peterson, J. A. (1996) J. Biol. Chem 271, 22663-22671.

13. Graham-Lorence, S., Truan, G., Peterson, J. A., Flack, J. R., WeL S., Helvig, C., Capdevilla, J. H. (1997) J. Biol. Chem 272, 1127-1135.

14. Li, H., Poulos, T. L. (1997) Nat. Structural Biol., 4, 140-146.

15. Ravichandran, K G., Sekhar, S., Boddupalli, S., Hasemann, C. A., Peterson, J. A., Deisenhofer, 1 (1993) Science 261, 731-736.

16. Modi S., Sutcliffe, M. J., Primrose, W. U., Lian, L.- Y., Roberts, G. C. K (1996) Nat. Structural Biol. 3, 414-417.

17. Oliver, C.F., ModL S., Primrose, W.U., Lian, L.Y. and Roberts, G.C.K (1997) Biochem. J. 327, 537-544.

18. Guengerich, F.G. (1991) J. Biol. Chem 266,10019-10022.

19. Cherry, J. R., Lamsa, M. H., Schneider, P., Vind, J., Svendsen, A-, Jones, A., and Pedersen, A. H. (1999) Nature Biotechnology 17, 379-384.

20. Schwaneberg, U., Schmidt-Dannert, C., Schmitt, J., and Schmid, R. D. (1999) Anal. Biochem. 269, 359-366.

21. Schwaneberg, U, Sprauer, AL, Schmidt-Dannert, C., and Schmid, R. D. J of Chromatogr. A, in press.

22. Oliver, C.F., Modi, S., Sutcliffe, M.J., Primrose, W.U., Lian, L.Y. and Roberts, G.C.K (1997) Biochemistry 36, 1567-1572.

23. Hart, S., Koch, KR., and Woods, D.R. (1992) J Gen. Microbiol. 138, 211-216

24. Murdock, D., Ensley, B.D., Serdar, C. and Thalen, M. (1993) Bio/Technology 11, 381-385.

25. O'Connor, ICE., Dobson, A-W. and Hartmans, S. (1997) Appl. Environ. Microbiol. 63, 4287-4291.

26. Eaton, R. W. and Chapman, P. J. (1995) J Bacteriol. 177, 6983-6988.